PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 : C12N 15/10, C12Q 1/68 // C12N 9/36, 15/56 — A1

(11) International Publication Number: **WO 94/12632**

(43) International Publication Date: 9 June 1994 (09.06.94)

(21) International Application Number: PCT/GB93/02436

(22) International Filing Date: 26 November 1993 (26.11.93)

(30) Priority Data:

9224882.2	27 November 1992 (27.11.92)	GB
9307132.2	6 April 1993 (06.04.93)	GB

(81) Designated States: JP, US, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.

(71) Applicant *(for all designated States except US)*: UNIVERSITY COLLEGE LONDON [GB/GB]; Gower Street, London WC1E 6BT (GB).

(72) Inventors; and

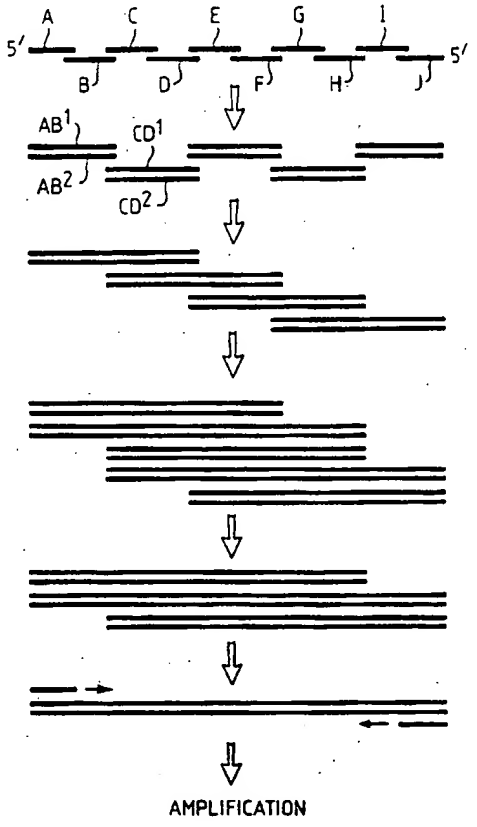
(75) Inventors/Applicants *(for US only)*: PRODROMOU, Chrisostomos [GB/GB]; 18 Dewsgreen, Vange, Basildon, Essex SS16 4NT (GB). PEARL, Laurence, Harris [GB/GB]; 24 Hurst Street, Herne Hill, London SE24 0EG (GB).

(74) Agent: PERCY, Richard, Keith; Patents Dept., British Technology Group Limited, 101 Newington Causeway, London SE1 6BU (GB).

(54) Title: IMPROVEMENTS IN NUCLEIC ACID SYNTHESIS BY PCR

(57) Abstract

A method of synthesising nucleic acid by a polymerase chain reaction carried out on a series of overlapping primers which span the entire length of the nucleic acid to be synthesised, comprising two 5'-end outermost primers and inner primers therebetween, each of the outermost primers being present initially in a sufficient excess over each of the inner primers to select for chain-extension of the outermost primers and their chain-extension products, over inner primers and their chain-extension products, characterised in that: at least 6 primers are used; and the reaction is carried out continuously, without any intermediate step of separating nucleic acid molecules of shorter length than required for the total synthesis by PCR of substantially the full length nucleic acid.



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	**GB**	United Kingdom	**MR**	Mauritania
AU	Australia	**GE**	Georgia	**MW**	Malawi
BB	Barbados	**GN**	Guinea	**NE**	Niger
BE	Belgium	**GR**	Greece	**NL**	Netherlands
BF	Burkina Faso	**HU**	Hungary	**NO**	Norway
BG	Bulgaria	**IE**	Ireland	**NZ**	New Zealand
BJ	Benin	**IT**	Italy	**PL**	Poland
BR	Brazil	**JP**	Japan	**PT**	Portugal
BY	Belarus	**KE**	Kenya	**RO**	Romania
CA	Canada	**KG**	Kyrgystan	**RU**	Russian Federation
CF	Central African Republic	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CG	Congo			**SE**	Sweden
CH	Switzerland	**KR**	Republic of Korea	**SI**	Slovenia
CI	Côte d'Ivoire	**KZ**	Kazakhstan	**SK**	Slovakia
CM	Cameroon	**LI**	Liechtenstein	**SN**	Senegal
CN	China	**LK**	Sri Lanka	**TD**	Chad
CS	Czechoslovakia	**LU**	Luxembourg	**TG**	Togo
CZ	Czech Republic	**LV**	Latvia	**TJ**	Tajikistan
DE	Germany	**MC**	Monaco	**TT**	Trinidad and Tobago
DK	Denmark	**MD**	Republic of Moldova	**UA**	Ukraine
ES	Spain	**MG**	Madagascar	**US**	United States of America
FI	Finland	**ML**	Mali	**UZ**	Uzbekistan
FR	France	**MN**	Mongolia	**VN**	Viet Nam
GA	Gabon				

IMPROVEMENTS IN NUCLEIC ACID SYNTHESIS BY PCR

Background of the invention

1. Field of the invention

This invention is the field of recombinant nucleic acid technology and more particularly relates to the polymerase chain reaction (PCR).

2. Description of the related art

The synthesis of even relatively short genes has generally been difficult and time-consuming. Both DNA strands are completely synthesised as short overlapping oligonucleotides which are phosphorylated, annealed and ligated to generate the full-length product (Edge et al., 1981; Bell et al., 1988; Ferretti et al., 1986). The cost of the synthesis can be reduced by synthesising oligonucleotides representing the partial sequence of each strand and the gaps in the annealed product 'filled in' using DNA polymerase prior to ligation (Rink et al., 1984). In practice, both these methods give a low yield of the full-length product and require amplification by cloning before any further manipulation of the synthesised gene.

Recently, a PCR procedure has been described in which a 234 base oligonucleotide was chemically synthesised and primers used to amplify any full-length molecules that resulted from the chemical synthesis (Barnet and Erfle, 1990). However, the length of gene synthesisable is still relatively short. Sandhu et al. (1992) have described a method of synthesising DNA by PCR carried out on a series of 4 overlapping DNA molecules of alternating complementary strands which span the entire length of the DNA to be synthesised. Each of the 5'-end outermost DNA molecules (H1, H4) is present initially in a sufficient excess over each of the inner DNA molecules (H2, H3) to select for chain-extension of the outermost DNA molecules and their chain-extension products over inner DNA molecules and their chain-extension products. In a detailed example, H1 to H4 are oligonucleotides of alternating complementary strands of lengths 86, 78, 63 and 17, with overlaps

of 12-17 bp., and are used to make ds DNA 220 bp. long. The authors also suggest using four 100-mers with 15 bp overlaps to produce ds DNA of 355 bp. They further suggest that to construct a longer synthetic gene two 355-mers overlapping by 15 bp should be made in separate reactions and then subjected to a fresh PCR to yield a 695-mer.

The Sandhu et al. method requires neither phosphorylation nor ligation and is relatively cheap as only oligonucleotides representing the partial sequence of each strand need to be chemically synthesised, as in the 'DNA-polymerase filling-in' method. However, it is a problem that the Sandhu et al. method has a practical upper limit of about 355 bp for the DNA to be synthesised, this upper limit being dictated (a) by the fact that it is at present difficult to synthesise oligonucleotides of lengths exceeding 100 bases in reasonable yield and (b) by the need for 15 bp of overlap [(4 x 100) - (3 x 15) = 355] between each of the four oligonucleotides employed. While longer genes could be synthesised, this requires a second reaction.

Summary of the invention

It has now been found that Sandhu et al. method can be improved to enable longer nucleic acid molecules to be synthesised more easily. According to the present invention, the Sandhu et al. method is varied and improved by using at least 6 overlapping starting nucleic acid molecules, each of which can correspond to either strand of the nucleic acid molecule to be synthesised, and by carrying out the reaction continuously, without any intermediate step of separating nucleic acid molecules of shorter lengths than required for the total synthesis by PCR of substantially the full length nucleic acid. Preferably the full length nucleic acid thus synthesised exceeds 400, most preferably exceeds 500 bases or base-pairs in length.

In the following description, the starting nucleic acid molecules will be referred to as primers. Ordinarily they are oligonucleotides, a term which signifies that they are of relatively short length, for example less than 120 bases long, compared with the length of nucleic acid to be synthesised.

The primers are mixed and subjected to PCR, so that their 3'-ends are extended to give longer double-stranded products. This is repeated for the double-stranded products until the full-sized nucleic acid is obtained. The full-length product is subsequently amplified by the 5'-end outermost oligonucleotides of each strand, which are present at a higher concentration than the inner ones and act as primers for the amplification.

Brief description of the drawings

Fig.1 is a schematic diagram of an embodiment of the method of the invention, explaining the chain-extension occurring;

Fig. 2 shows the oligonucleotides used in the synthesis of 522 bp long DNA encoding the lysozyme gene;

Fig. 3 is a photograph of a stained agarose gel showing that a full length gene of correct relative molecular mass has been synthesised by the method of the invention;

Fig. 4 is a schematic diagram which shows an embodiment of the invention in which the method is applied to plasmid DNA, enabling a gene to be synthesised directly into a plasmid.

Fig. 5 is a schematic diagram similar to that of Fig. 1, but showing another embodimentof the method of the invention using a different arrangement of primers; and

Fig. 6 is another schematic diagram, in briefer form, showing a further arrangement of primers.

Description of the preferred embodiments

A summary of particularly preferred features of the invention can be found in claims 2 onwards, the text of which should be regarded as part of the description also.

In the ordinary application of the invention the nucleic acid to be synthesised will be DNA and most usually a gene, i.e. a length of nucleic acid which comprises a region coding for a protein, preferably together with any associated sequences required or preferred for control of or aiding transcription or expression, e.g. a promoter region. However, the function of the nucleic acid to be synthesised is immaterial to the invention. It could be used, for example, to synthesise the intron DNA of a genome, which may be useful in gene therapy.

There are no lower limits or known upper limits on the lengths of the nucleic acid to be synthesised. However, it is intended that the invention should be applied mainly to be synthesis of long lengths, e.g. exceeding 400, preferably exceeding 500 bases or base-pairs, since it is in such a context that its advantages are best employed. It will readily be possible to synthesise DNA of up to 1 kb(p) and even longer e.g. 2 or 5 kb(p). The upper limit will be dictated currently only by deviations of the conventional PCR technology from perfection, but currently would be unlikely to exceed 10^6 kb(p).

The initial product of the PCR method of invention is double-stranded nucleic acid, but if desired this could be made single-stranded (subject always to stability, avoiding looping back on itseif etc.).

Primers will normally be oligonucleotides, which will rarely exceed 100 bases in length. The 5'-end outermost primers can be provided by double-stranded nucleic acid having a protruding "sticky end" of single-stranded nucleic acid which has a 3'-terminus free. This single-stranded-nucleic acid serves as the primer, whereby the chain-extended molecule will contain the double-stranded-nucleic acid. This embodiment is illustrated later by reference to Fig. 4.

Although the invention is mainly of interest for PCR using wholly DNA primer and a DNA polymerase for chain extension, it is possible to apply it to reverse transcriptase PCT (RT-PCR), using mixed RNA and DNA primers, the outermost primers being of DNA, or even to wholly RNA primers using an RNA polymerase.

The operation of the invention is most easily seen from the simplified diagram of Fig. 1. The figure shows six stages of one embodiment of the method, although these divisions are purely notional since synthesis is continuous, "in one pot". The first stage is the beginning of a reaction involving 10 overlapping oligonucleotides. The products at stage 2 have undergone one cycle, in which, for example, the left-hand, inner oligonucleotide marked "B" is extended from its 3'-end using the 5'-end outermost oligonucleotide "A" as template and likewise "A"

is extended from its 3'-end using the inner oligonucleotide "B" as template. A hybrid (AB) is thus formed and separated into two complementary strands (AB^1, AB^2). Other hybrids (CD etc.) are also formed and, after the denaturing step, separated into single strands (CD^1, CD^2 etc.), along all the length of the nucleic acid to be synthesised. Because the 5'-end outermost oligonucleotides (A, J) are in great excess over the inner oligonucleotides (B to I), the next cycle will result in increased amounts of AB hybrid, since primer A will be extended along the length of AB^2 preferentially to primer C being extended along the length of CD^2 and so on. (This preponderance is not shown in Figure 1). The nucleic acid synthesis proceeds in each cycle as shown until finally full length DNA is produced. This full length nucleic acid will also be amplified preferentially to any shorter nucleic acid requiring priming from an inner primer and so the process is self-selective for the full length product. Of course, the scheme shown in Fig. 1 is arbitrary in the sense that the intermediate hybrids formed are mixtures of (initially) AB, BC, CD, DE, EF, FG, GH, HI and IJ, followed by ABC, BCD, CDE and so on.

Complete synthesis of the nucleic acid results from the mutual extension of two halves which themselves result from the mutual extension of two halves etc. This is functionally identical to the technique of recursion in computer programming and we therefore describe this embodiment as 'Recursive PCR'.

A sophisticated development of the above embodiment of the invention is shown in Fig. 4. The outermost DNA molecules are plasmid DNA which has been cut. For ease of illustration, the outermost (A, J) and inner (B to I) primers are labelled correspondingly to Fig. 1. A and J are joined together as duplex (ds) DNA in the plasmid in the area between two sites, at which 3' and 5'-ends are marked. The 3'-end overlaps and is complementary to the 5'-end of B, while the 3'-end of J overlaps and is complementary to the 5'-end of I. The resultant full length molecules will be circular, and capable of transforming bacteria without further treatment.

It is not necessary that successive inner primers be of alternate strands of the nucleic acid to be synthesised. So long as they are overlapping they can be of either or both strands. In the embodiment shown in Figure 5, for example, all the inner primers (L to R) are of the same strand. Of course, the outermost primers (K,S) have to be of different strands. Analogously to Figure 1, the reaction can be regarded as proceeding from the left-hand end via initial hybrids KLM^1, KLM^2, $KLMO^1$, $KLMO^2$ ... to the full length sequence $KLMNOPQRS^1$, $KLMNOPQRS^2$. After strand separation the primers K and S operate on the entire strand as template, priming in both directions, as for Fig. 1. Nine primers in all are shown in Figure 5, but it is immaterial whether the number is odd or even, and fewer or more primers than nine could be used according to the length of nucleic acid to be synthesised.

In Figure 6, which shows another possible arrangement, involving a total of 13 primers, one of the middle internal primers is of different strand from the others. The arbitrary stages, analogous to a mixture of those shown in Figures 1 and 5, have been omitted from Figure 6 for brevity.

The outermost primers are both 5'-ended. Any 3'-ended outermost primer having the sequence of a part of the nucleic acid to be synthesised lying beyond the outermost 5'-ends of the primers would simply not be copied and the full length DNA would extend only between the two 5'-outermost ends of the series of primers. In other words the nucleic acid to be synthesised is <u>by definition</u>, that which lies between the 5'-end outermost primers. However, the outermost primers may be double-stranded and this embodiment has some advantage, because polymerases can chew back the ends of the single-stranded primers. For instance, in one experiment in which the invention was applied to synthesising the 1044 bp ornithine carboxylase gene, using 18 primers, a 71-base outermost primer was chewed back by 8-bases and then began to prime on to innder primer No. 14 with which the chewed back sequence had 9 bp of perfect homology.

The full range of ratios of the 5'-end outermost to inner DNA primers is not capable of definition by sharply defined numerical limits. On the one hand, a high ratio is very favourable because it will increase selection for full length products and molar ratios within the range 50-500:1, preferably 100-500:1, will be of particular importance. While higher ratios are usable, they will often imply such a low amount of inner primer that the PCR will not readily proceed. Such a consideration will be well understood by those skilled in the art. On the other hand, a molar excess of (say) only 10 times might well not work, depending, however, on the other variables such as lengths of overlaps in the series of primers, lengths of gaps to be filled in, the number and length of the primers, and the proof-reading accuracy of the polymerase or other chain-extending enzyme employed.

Successive primers (which mutually serve as "primer and template") will normally overlap, generally by a length predominantly within the range 15-25 bp, preferably 17-20 bp. This applies to the starting primers and to primers formed by the initial hybrids, e.g. KLM[1] in Figure 5 which primes for further nucleic acid synthesis on the template provided by O. Generally stated, usual "rules" for arranging priming conditions in the PCR apply. One such set of "rules" (for all DNA PCR) is to arrange for the overlapped region to have a melting temperature of 45-72°C, preferably 50-60°C, and most preferably 52-56°C. Melting temperature is roughly calculable by allowing 4°C for every G-C pairing and 2°C for every A-T pairing in the prospective duplex of the "overlapped region".

Preferably the proportion of G-C pairs in the overlapped region should not exceed 67% (two thirds) and it is preferred that it does not exceed 55% with less than 45% even more preferred.

Normally, the series of primers will not be abutting within a single prospective strand. That is, there will be gaps to be filled in by chain-extension. However, this rule is not critical and it is possible to use abutting molecules. Thus, for instance,

in Figure 1, instead of primers E and G of the first strand, each overlapping F of the second strand, primers EX and G could be used instead, where X represents the sequence between the 3'-end of E and the 5'-end of G in Figure 1.

The lengths of the gaps to be filled in can be varied widely, but assuming arbitrarily a primer length of from 50 to 100 and an overlap of 17-20 at each end thereof, they would usually be in the range 10 to 66. The preferred length of primer is arbitrary since there is a trade-off between the ease of oligonucleotide synthesis and the ease of performing this reaction, but will usually be in the 50 to 100 range, preferably predominantly 70 to 90.

It is preferred not to use Taq polymerase as the chain-extending enzyme in the PCR, because there are now available enzymes of superior proof-reading ability which are preferred. They include Vent, Deep Vent and Pfu DNA polymerases.

Otherwise the conditions are generally as for amplifying PCRs. The primers should not contain any undue complementarity with other primers, to avoid unwanted primer-primer interactions. If the DNA to be synthesised has any tendency for self-complementarity leading to looping, regions with this property should be split across two overlapping primers.

In general, cycling temperatures and times in the chain-extension, annealing and denaturing steps can be those conventionally used. It is preferred to anneal at a temperature not more than 10°C below the calculated melting temperature of the prospective duplex. The final chain extension step is preferably carried out at elevated temperatures, up to 72°C, for (say) 10 minutes, to ensure completion of the reaction.

The following Example illustrates the invention.

Example

MATERIALS AND METHODS

Ten oligonucleotide primers, representing a synthetic human-lysozyme gene sequence were synthesized using ß-cyanoethyl phosphoramidite chemistry and purified on polyacrylamide gels by British Biotechnology Limited Figure 2 shows these primers in the

conventional left to right = 5'-end to 3'-end manner. Their sequences are shown in SEQ ID NO: 1 to 10 hereinafter. The oligonucleotide primers were 54-86 bases long to minimise errors and obtain good yields in the chemical synthesis. The overlapping regions 17-20 bp in length, were designed to give melting temperatures mainly in the range 52-56°C, but in some instances up to 60°C to ensure good primer specificity and a similar degree of annealing at all overlaps in the initial cycles of the PCR reaction. A computer search of both strands of the entire sequence with the sequences of each of the overlap regions showed them to be unique and unlikely to give non-specific priming. Subsequent insertion of the synthesised gene into expression vectors, was facilitated by the incorporation in the outermost oligonucleotide primers, of an EcoRI site at the 5' end and a HindIII site at the 3' end of the gene. A Shine-Dalgarno sequence and the TØ transcription terminator (from gene 10B of bacteriophage T7) were incorporated into the outermost 5' and 3' oligonucleotide primers, respectively, for the efficient expression of the gene in Escherichia coli. Sequences such as the TØ transcription terminator can potentially inhibit the PCR reaction by formation of strong stem loops. This was prevented by incorporating the two halves of the stem loop in separate overlapping oligonucleotide primers (see Figure 2). In addition to the EcoRI and HindIII sites, fourteen additional unique restriction sites were incorporated into the gene sequence to aid subsequent sub-cloning and mutagenesis (see Figure 2). The coding region was designed to contain the most frequently used codons found in highly expressed genes in Saccharomyces cerevisiae, an intended expression host, except where this conflicted with the unique restriction sites. As the codon usage of E. coli is not as biased as that of S. cerevisiae, the use of optimal S. cerevisiae codon usage will still allow efficient expression in E. coli.

The PCR reaction (30 cycles of: 2 min at 95°C, 2 min at 56°C and 1 min at 72°C, with a final 10 min at 72°C) was performed in a Techne PHC-2 thermal cycler. The annealing temperature of 56°C was chosen to minimise non-specific priming by the overlapping

oligonucleotides, (see above). The reaction contained 20 - 30 pmoles of the outermost 5' oligonucleotides of each strand, 2 - 3, 0.2 - 0.3, or 0.02 - 0.03, pmoles of the internal oligonucleotides, 10 µl 10 X Vent-buffer, 4 U Vent DNA-polymerase (New England Biolabs), and 6 µl 100 mM MgSO4, in a total reaction volume of 100 µl overlaid with 50 µl paraffin oil to prevent evaporation. The full-sized PCR product (522 bp) was isolated from a 1.5% low-melting-point agarose gel, purified by sequential extraction with phenol, phenol/chloroform, chloroform and ether, ethanol precipitated and finally 200 ng of the purified PCR product digested in a total volume of 20 µl containing 1 U EcoRI, 1 U HindIII (Northumbria Biochemicals Ltd.) and 2 µl 10 X buffer. The restriction enzymes were extracted with Strataclene resin (Stratagene). The DNA was ethanol-precipitated and ligated to 200 ng of similarly digested phagemid, pGEM3Zf(+) (Promega), in a total volume of 20 µl, containing 0.5 U T4 DNA ligase (Northumbria Biochemicals Ltd.) and 2 µl 10 X buffer. E. coli DH5a was transformed and plated onto LB plates containing ampicillin (100 mg/ml), X-gal (20 mg/ml) and IPTG (32 mg/ml). Double-stranded phagemid was sequenced by the dideoxy-chain-termination method (Sanger et al., 1977) using Sequenase (Tabor and Richardson, 1987).

RESULTS AND DISCUSSION

We have used the method of the invention to synthesise a human lysozyme gene of 522 bp using 10 overlapping oligonucleotides in a single reaction. To our knowledge, this is the largest double-stranded DNA molecule to be synthesised to date, in a single step.

Figure 3 shows a photograph of a stained agarose gel. Lanes 1-3 correspond to the three molar ratios of 5'-outermost: inner oligonucleotide primers used, i.e. 10:1, 100:1 and 1000:1 lane 4 shows markers. The lysozyme gene synthesised is arrowed. It is likely that the inner oligonucleotide primers, which can only be extended in the normal 5' to 3' direction, are inhibitory to the synthesis of the full-length product (Figure 3, lane 1). However, at internal oligonucleotide primer concentrations 100 times lower

than that of the outer oligonucleotide primers (2 - 3 pmoles/µl) efficient and selective synthesis of the full-length product is obtained (Lane 2). Compare the marker of relative molecular mass 514 base pairs in lane 4. At inner oligonucleotide primer concentrations 1000 times lower than those of the outermost oligonucleotide primers, no visible amplification was observed. This is ascribed to use of too small an amount of the inner oligonucleotide primers (lane 3).

REFERENCES

Barnet,R.W. and Erfle,H. (1990) Nucleic Acids Res., 18, 3094.

Bell,L.D., Smith,J.C., Derbyshire,R., Finlay,M., Johnson,I., Gilbert,R., Slocombe,P., Cook,E., Richards,H., Clissold,P., Meredith,D., Powell-Jones,C.H., Dawson,K.M., Carter,B.L. and McCullagh,K.G. (1988) Gene, 63, 155-163.

Edge,M.D., Greene,A.R., Heathcliffe,G.R., Meacock,P.A., Schuch,W., Scanlon, D.B., Atkinson,T.C., Newton,C.R. and Markham,A.F. (1981) Nature, 292, 756-762.

Ferretti,L., Karnik,S.S., Khorana,H.G., Nassal,M. and Oprian,D.D. (1986) Proc.Natl.Acad.Sci.USA, 83, 599-603.

Mead,D.A., Pey,N.K., Herrnstadt,C., Marcil,R.A. and Smith,L.M. (1991) Biotechnology, 9, 657-663.

Rink,H., Manfred,L., Sieber,P. and Meyer,F. (1984) Nucleic Acids Res., 15, 6369-6387.

Sandhu,G.S., Aleff,R.A. and Kline,B.C. (1992) Biotechniques, 12, 14-16.

Sanger,F., Nicklen,S. and Coulsen,A.R. (1977) Proc.Natl.Acad.Sci.USA, 74, 5463-5467.

Tabor,S. and Richardson,C.C. (1987) Proc.Natl.Acad.Sci.USA, 84,4767-4771.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:
(A) NAME: UNIVERSITY COLLEGE LONDON
(B) STREET: Gower Street
(C) CITY: London
(E) COUNTRY: GB
(F) POSTAL CODE (ZIP): WC1E 6BT

(A) NAME: PEARL, Laurence Harris
(B) STREET: 24 Hurst Street, Herne Hill
(C) CITY: London
(E) COUNTRY: United Kingdom
(F) POSTAL CODE (ZIP): SE24 0EG

(A) NAME: PRODROMOU, Chrisostomos
(B) STREET: 18 Dewsgreen, Vange
(C) CITY: Basildon
(D) STATE: Essex
(E) COUNTRY: United Kingdom
(F) POSTAL CODE (ZIP): SS16 4NT

(ii) TITLE OF INVENTION: IMPROVEMENTS IN NUCLEIC ACID SYNTHESIS BY PCR

(iii) NUMBER OF SEQUENCES: 10

(iv) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

(v) CURRENT APPLICATION DATA:
APPLICATION NUMBER: WO Not yet assigned
(vi) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: GB 9224882.2
(B) FILING DATE: 27-NOV-1992

(vi) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: GB 9307132.2
(B) FILING DATE: 06-APR-1993

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 77 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Oligo primer for synthesis of lysozyme gene

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

GCCGGAATTC AAATAATTTG TTTAACTTTA AGGAGATCTA CATATGAAGG TTTTCGAAAG 60

ATGTGAGCTC GCTAGAA 77

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 81 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(v) FRAGMENT TYPE: N-terminal

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Oligo primer for synthesis of lysozyme gene

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

CCAAACACAT CCAGTTAGCC AAAGAGATAC CACGGTACCC GTCCATACCC AATCTCTTCA 60

AGGTTCTAGC GAGCTCACAT C 81

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 86 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Oligo primer for synthesis of lysozyme gene

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

GCTAACTGGA TGTGTTTGGC TAAGTGGGAA TCTGGTTACA ACACTCGAGC TACCAACTAC 60

AACGCTGGTG ACAGGTCGAC CGACTA 86

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 81 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Oligo primer for synthesis of lysozyme gene

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

TTAACGGCGC CTGGGGTCTT ACCGTCGTTA CACCAGTATC TAGAATTAAT TTGGAAGATA 60

CCGTAGTCGG TCGACCTGTC A 81

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 68 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iii) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Oligo primer for synthesis of lysozyme gene

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

GACCCCAGGC GCCGTTAACG CTTGTCACTT GTCTTGTTCT GCTTTGCTGC AGGACAACAT 60

CGCTGACG 68

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 58 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Oligo primer for synthesis of lysozyme gene

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

CCTTGTGGGT CTCTAACAAC GCGTTTAGCA CAAGCAACAG CGTCAGCGAT GTTGTCCT 58

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 55 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Oligo primer for synthesis of lysozyme gene

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

TGTTAGAGAC CCACAAGGTA TCAGAGCTTG GGTTGCTTGG AGAAATCGAT GTCAA 55

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 68 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Oligo primer for synthesis of lysozyme gene

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

TCCTTACTAA ACACCACAAC CTTGAACGTA TTGTCTGACG TCTCTGTTTT GACATCGATT 60

TCTCCAAG 68

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 60 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Oligo primer for synthesis of lysozyme gene

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

GTTGTGGTGT TTAGTAAGGA TCCTCTAGAA TAACCCCTTG GGGCCTCTAA ACGGGTCTTG 60

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 54 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Oligo primer for synthesis of lysozyme gene

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

CCGATGCCGA AGCTTCCTCC ATGGAGCAAA AACCCCTCAA GACCCGTTTA GAGG 54

CLAIMS

1. A method of synthesising nucleic acid by a polymerase chain reaction carried out on a series of overlapping primers which span the entire length of the nucleic acid to be synthesised, comprising two 5'-end outermost primers and inner primers therebetween, each of the outermost primers being present initially in a sufficient excess over each of the inner primers to select for chain-extension of the outermost primers and their chain-extension products, over inner primers and their chain-extension products, characterised in that:

- at least 6 primers are used; and
- the reaction is carried out continuously, without any intermediate step of separating nucleic acid molecules of shorter length than required for the total synthesis by PCR of substantially the full length nucleic acid.

2. A method according to claim 1, characterised in that the initial molar ratio of 5' end outermost to inner primers is at least 100:1

3. A method according to claim 2, characterised in that the ratio is from 100:1 to 500:1.

4. A method according to any preceding claim characterised in that nucleic acid of length exceeding 400 bases or base-pairs is synthesised.

5. A method according to claim 4 characterised in that nucleic acid of length 500 to 5000 bases or base-pairs is synthesised.

6. A method according to any preceding claim characterised in that the primers are of alternating complementary strands.

7. A method according to any preceding claim, characterised in that the primers are wholly of DNA.

8. A method according to claim 7 characterised in that the series of primers provides a prospective ds DNA sequence having gaps within each prospective strand to be filled in by chain extension and in that these gaps are from 10 to 66 bases in length.

9. A method according to claim 7 or 8, characterised in that all the primers are arranged to have or impart an overlap of from 17-20 bp calculated to impart a melting temperature to the hybrid in the area of overlap of from 50 to 60°C, said melting temperature being calculated by allowing 4°C for each G-C pairing and 2°C for each A-T pairing, provided that G-C pairings do not exceed 67%.

10. A method according to claim 7, 8, or 9, characterised in that it is carried out using a chain-extending enzyme of superior proof-reading ability to Taq polymerase.

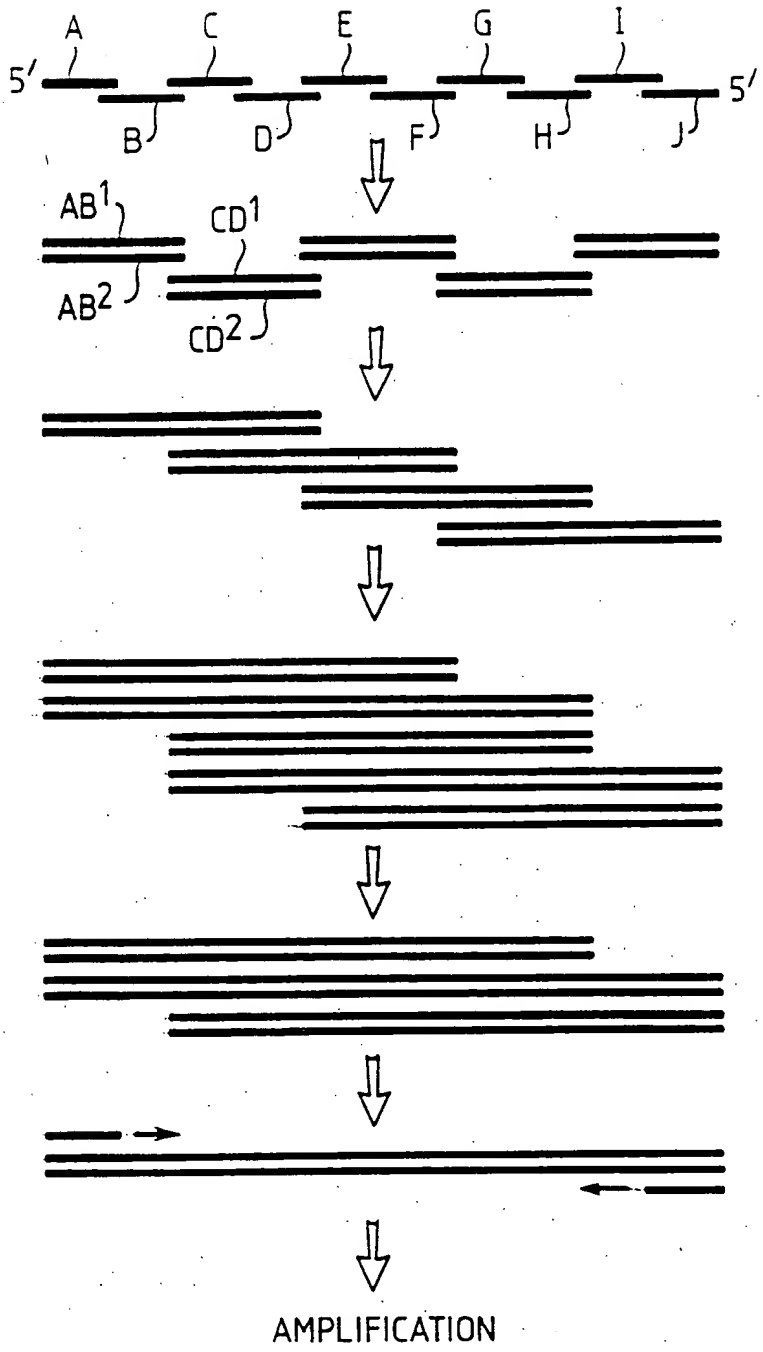


Fig. 1

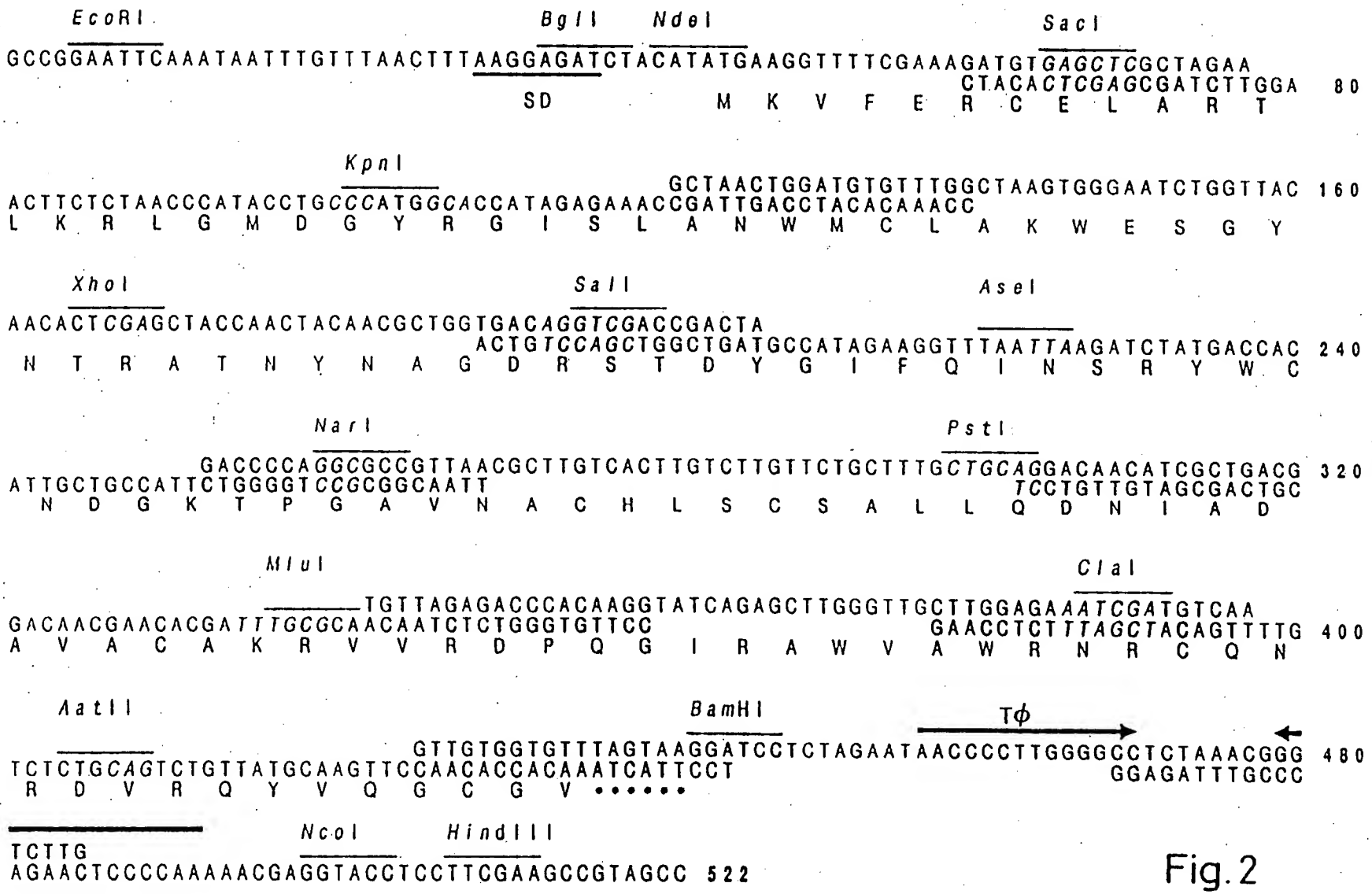

EcoRI
BglI
NdeI
SacI
SD
M K V F E R C E L A R T
KpnI
L K R L G M D G Y R G I S L A N W M C L A K W E S G Y
XhoI
SalI
AseI
N T R A T N Y N A G D R S T D Y G I F Q I N S R Y W C
NarI
PstI
N D G K T P G A V N A C H L S C S A L L Q D N I A D
MluI
ClaI
A V A C A K R V V R D P Q G I R A W V A W R N R C Q N
AatII
BamHI
Tϕ
R D V R Q Y V Q G C G V ••••••
NcoI
HindIII
TCTTG
AGAACTCCCCAAAAACGAGGTACCTCCTTCGAAGCCGTAGCC 522

Fig. 2

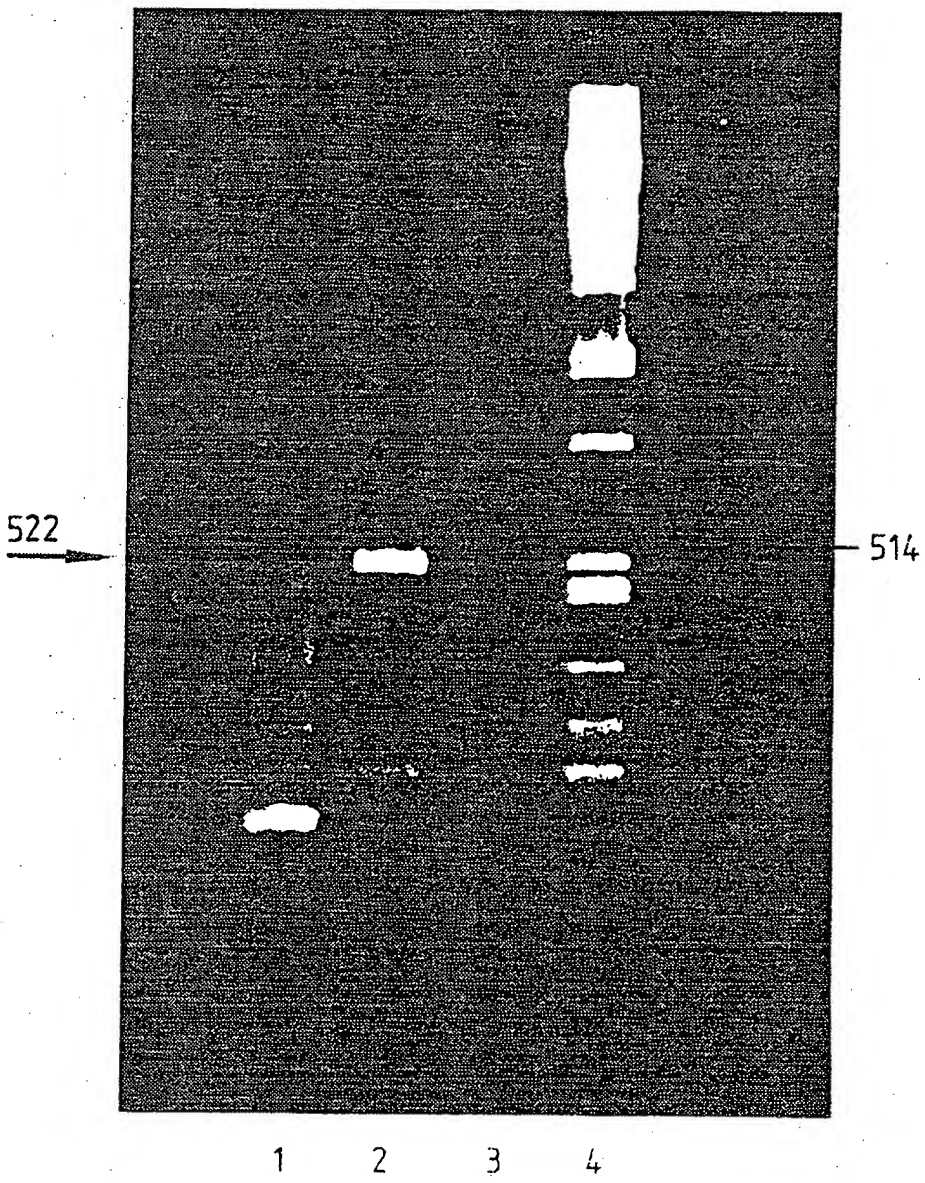


Fig. 3

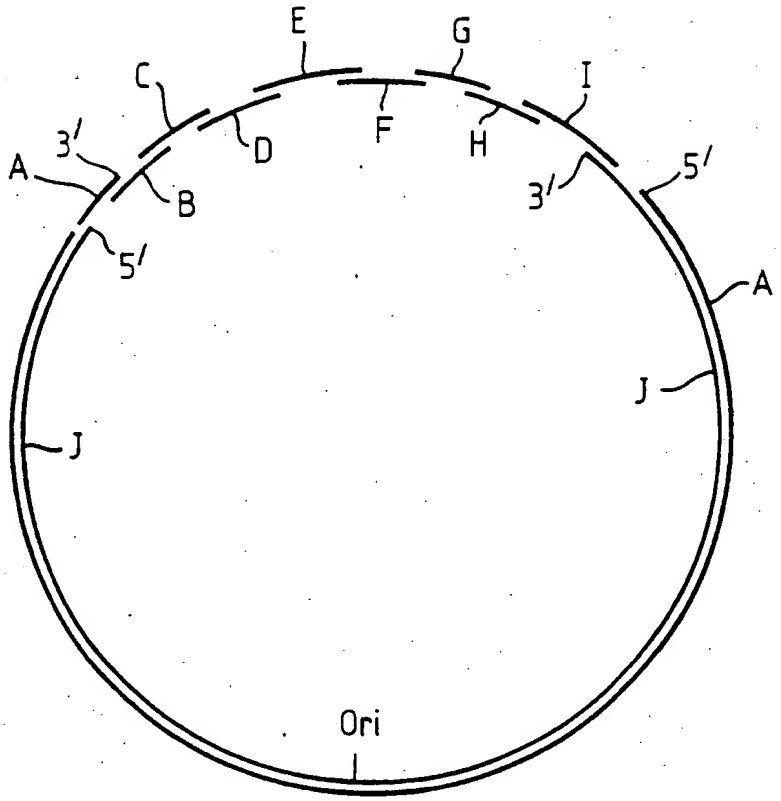


Fig. 4

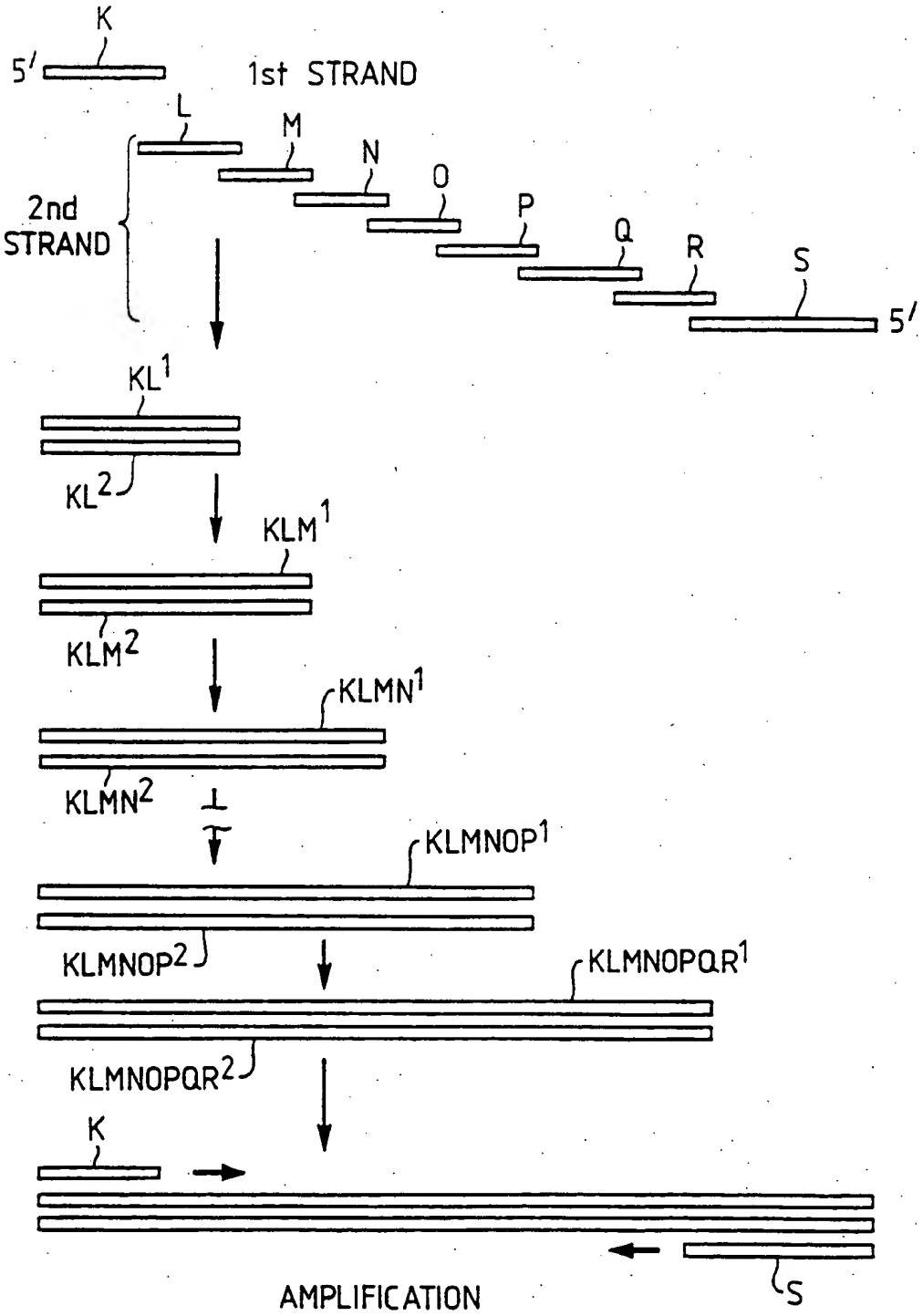

K
5'
1st STRAND
L
M
N
O
P
Q
R
S
5'
2nd STRAND
KL1
KL2
KLM1
KLM2
KLMN1
KLMN2
KLMNOP1
KLMNOP2
KLMNOPQR1
KLMNOPQR2
K
S
AMPLIFICATION

Fig. 5

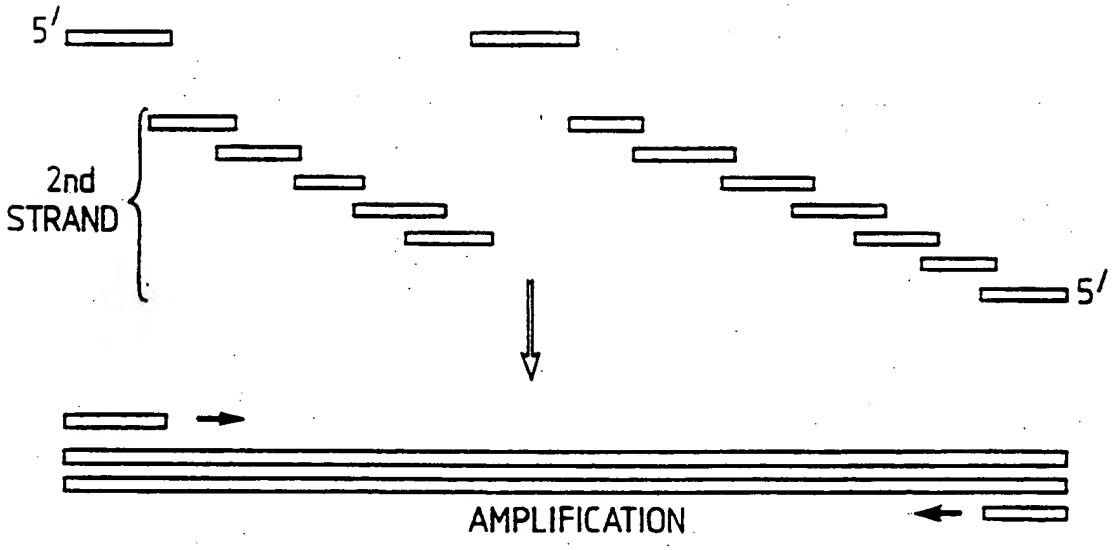


Fig. 6

INTERNATIONAL SEARCH REPORT

International Application No
PCT/GB 93/02436

A. CLASSIFICATION OF SUBJECT MATTER
IPC 5 C12N15/10 C12Q1/68 //C12N9/36,C12N15/56

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 5 C12N C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	BIOTECHNIQUES vol. 12, no. 1 , January 1992 , EATON PUBL. CO., MA,US; pages 14 - 16 G.S. SANDHU ET AL. 'Dual asymmetric PCR: One-step construction of synthetic genes' cited in the application the whole document ---	1-7
X	CURRENT PROTOCOLS IN MOLECULAR BIOLOGY 1989 , WILEY & SONS, INC.,US; pages 8.2.8 - 8.2.13 'Gene synthesis: Assembly of target sequences using mutually priming long oligonucleotides' see page 8.2.10, left column, line 1 - page 8.2.13, right column, line 3; figure 8.2.5 ---	1-7

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
17 February 1994	04 -03- 1994

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax (+31-70) 340-3016	Hornig, H

Form PCT/ISA/210 (second sheet) (July 1992)

International Application No
PCT/GB 93/02436

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	EP,A,0 385 410 (CANON KABUSHIKI KAISHA) 5 September 1990 see page 3, line 29 - line 47; figures 2,3 see page 5, line 21 - page 7, line 35	1-7
Y	EP,A,0 316 018 (CETUS CORPORATION) 17 May 1989 see page 2, line 1 - line 39; claims 1-7; figure 1	1-10
Y	GENE vol. 77, no. 1 , 1989 , ELSEVIER PUBLISHERS, N.Y., U.S.; pages 61 - 68 R.M. HORTON ET AL. 'Engeneering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension' see page 61, line 1 - line 10; figure 2	1-10
A	GENE vol. 76, no. 1 , 1989 , ELSEVIER PUBLISHERS, N.Y., U.S.; pages 153 - 160 M.A. WOSNICK ET AL. 'Total chemical synthesis and expression in Escherichia coli of a maize glutathione-transferase (GST) gene' see page 154, right column, line 1 - page 157, right column, line 25; figures 1,2	1-10
P,X	PROTEIN ENGENEERING vol. 5, no. 8 , December 1992 , IRL PRESS OXFORD, ENGL. pages 827 - 829 C. PRODROMOU AND L.H. PEARL 'Precursive PCR: a novel technique for total gene synthesis' the whole document	1-10

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/GB 93/02436

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
EP-A-0385410	05-09-90	JP-A-	3210195	13-09-91
		JP-A-	3232489	16-10-91
		JP-A-	3232490	16-10-91
		JP-A-	3127999	31-05-91
		JP-A-	3007583	14-01-91
EP-A-0316018	17-05-89	EP-A-	0196762	08-10-86
		EP-A-	0314202	03-05-89
		JP-A-	61293916	24-12-86
		US-A-	5079163	07-01-92